CLAIMS

I Claim:

1. A method of recognizing a received phoneme using a stored plurality of phoneme classes, each of the plurality of phoneme classes comprising class phonemes, the method comprising:

(A) training the class phonemes, the training comprising, for each class phoneme:

(1) determining a phoneme vector as a time-frequency representation of the class phoneme;

(2) dividing the phoneme vector into phoneme segments;

(3) assigning each phoneme segment into a plurality of phoneme parameters;

(4) expanding each phoneme segment and plurality of phoneme parameters into an expanded stored-phoneme vector with expanded vector parameters;

(5) transforming the expanded stored-phoneme vector into an orthogonal form using singular-value decomposition wherein: $[x_1 x_2 \ldots x_m] = [u_1 u_2 \ldots u_m] \Lambda V^t$, where x_k is a k^{th} acoustic vector for a corresponding stored phoneme, u_k is the corresponding orthogonal vector and Λ and V are diagonal and unitary matrices, respectively; and

(B) recognizing the received phoneme by:

(1) receiving an analog acoustic signal;

(2) converting the analog acoustic signal into a digital signal;

(3) determining a received-signal vector as a time-frequency representation of the received digital signal;

(4) dividing the received-signal vector into received-signal segments;

(5) assigning each received-signal segment into a plurality of received-signal parameters;

(6) expanding each received-signal segment and plurality of received-signal parameters into an expanded received-signal vector;

(7) transforming the expanded received-signal vector into an orthogonal form using singular-value decomposition wherein: $[y_k] = [z_k]\,\Lambda V^t$, where y_k is a k^{th} acoustic vector for a corresponding received phoneme, z_k is the corresponding orthogonal vector and Λ and V are diagonal and unitary matrices, respectively;

(8) determining a first distance associated with the orthogonal form of the expanded received-signal vector and a second distance associated respectively with each orthogonal form of the expanded stored-phoneme vectors; and

(9) recognizing the received phoneme according to a comparison of the first distance with the second distance.

2. The method of claim 1, wherein transforming the expanded stored-phoneme vector into an orthogonal form using singular-value decomposition and wherein transforming the expanded received-signal vector into an orthogonal form using singular-value decomposition conforms the stored-phoneme vector and the expanded received-signal vector into a hypersphere having a center and a radius.

3. The method of claim 2, wherein determining a distance associated with the orthogonal form of the expanded received-signal vector and each orthogonal form of the expanded stored-phoneme vectors further comprises:

comparing a distance from the center of the hypersphere of the orthogonal form of the expanded received-signal vector with a distance from the center of the hypersphere for each orthogonal form of the expanded stored-phoneme vector.

4. The method of claim 3, wherein determining a distance associated with the orthogonal form of the expanded received-signal vector and each orthogonal form of the expanded stored-phoneme vectors further comprises:

determining a difference between the distance from the center of the hypersphere of the orthogonal form of the expanded received-signal vector and the distance from the center of the hypersphere for each orthogonal form of the expanded stored-phoneme vectors, wherein the expanded stored-phoneme vectors associated with m-shortest differences between the distance from the center of the hypersphere of the orthogonal form of the expanded received-signal vector and the distance from the center of the hypersphere for each orthogonal form of the expanded stored-phoneme vectors are recognized as most likely to be associated with the received phoneme.

5. The method of claim 1, wherein the orthogonal form of the expanded stored-phoneme vector and the expanded received-signal vector each have at least approximately 100 dimensions.

6. The method of claim 1, wherein each acoustic vector for a corresponding stored phoneme has a mean value removed.

7. The method of claim 6, wherein each acoustic vector for a corresponding received phoneme has a mean value removed.

8. The method of claim 1, wherein the phoneme vector determined as a time-frequency representation of the class phoneme is a representation of approximately 125 msec.

9. The method of claim 8, wherein the phoneme vector is divided into approximately 25 msec phoneme segments.

10. The method of claim 9, wherein each 25 msec phoneme segment is assigned approximately 32 phoneme parameters.

11. The method of claim 10, wherein each of the approximately 25 msec phoneme segments with 32 phoneme parameters is expanded into an expanded stored-phoneme vector with approximately 160 parameters.

12. The method of claim 11, wherein the received-signal vector determined as a time-frequency representation of the received digital signal is a representation of approximately 125 msec.

13. The method of claim 11, wherein the received-signal vector is divided into approximately 25 msec received-signal segments.

14. The method of claim 13, wherein each approximately 25 msec received-signal segment is assigned approximately 32 received-signal parameters.

15. The method of claim 14, wherein each of the approximately 25 msec received-signal segments with 32 received-signal parameters is expanded into an expanded received-signal vector with approximately 160 parameters.

16. A method of recognizing speech patterns, the method using stored phonemes, the method comprising:

converting each stored phoneme into n-dimensional space having a center;

sampling speech patterns to obtain at least one sampled phoneme;

converting each of the at least one sampled phonemes into the n-dimensional space; and

comparing a distance from the center of the n-dimensional space to the sampled phoneme with a distance from the center of the n-dimensional space to each of the phonemes of the converted plurality of phonemes.

17. The method of claim 16, wherein converting the stored phonemes comprises using singular-value decomposition.

18. The method of claim 16, further comprising storing the converted phonemes before sampling speech patterns.

19. The method of claim 16, wherein n equals at least 100.

20. The method of claim 16, wherein comparing the distance from the center of the n-dimensional space to the sampled phoneme with the distance from the center of the n-dimensional space to each of the converted phonemes further comprises: determining a difference between the distance from the center of the n-dimensional space to the sampled phoneme with the distance from the center of the n-dimensional space to each of the converted phonemes.

21. The method of claim 20, further comprising:

recognizing the sampled phoneme as the stored phoneme associated with the smallest difference between the distance from the center of the n-dimensional space to the sampled phoneme with the distance from the center of the n-dimensional space to each of the converted phonemes.

22. The method of claim 16, wherein the n-dimensional space is hyperspherical.

23. The method of claim 16, wherein converting the stored plurality of phonemes into n-dimensional space having a center further comprises:

assigning a stored-phoneme vector having approximately 160 parameters to each stored phoneme; and

transforming each stored-phoneme vector into the n-dimensional space having the center, wherein a probability density of the stored phonemes in the n-dimensional space is approximately spherical.

24. The method of claim 23, wherein converting each of the at least one sampled phonemes into the n-dimensional space further comprises:

assigning a sampled-phoneme vector having approximately 160 parameters to each sampled phoneme; and

transforming each sampled-phoneme vector into the n-dimensional space having the center, wherein a probability density of the stored phonemes in the n-dimensional space is approximately spherical.

25. A method of recognizing speech using a database of stored phonemes converted into n-dimensional space, the method comprising:

receiving a received phoneme;

converting the received phoneme to n-dimensional space;

comparing the received phoneme to each of the stored phonemes in n-dimensional space; and

recognizing the received phoneme according the comparison of the received phoneme to each of the stored phonemes.

26. The method of recognizing speech according to claim 25, wherein comparing the received phoneme to each of the stored phonemes in n-dimensional space further comprises:

comparing a first distance from a center of the n-dimensional space to a first point associated with the received phoneme with a second distance from the center of the n-dimensional space to a second point associated in turn with each of the stored phonemes.

27. The method of claim 26, wherein "n" is at least approximately 100.

28. The method of claim 26, wherein comparing the first distance with the second distance for each of the stored phonemes further comprises:

determining a difference between the first distance and the second distance for each stored phoneme.

29. The method of claim 28, wherein recognizing the received phoneme according the comparison of the received phoneme to each of the stored phonemes further comprises:

recognizing the received phoneme according to the stored phoneme associated with the smallest difference between the first distance and the second distance.

30. A system for recognizing phonemes, the system using a database of stored phonemes for comparison with received phonemes, the stored phonemes having been converted into n-dimensional space, the system comprising:

a recording element that receives a phoneme;

a computer that converts the received phoneme into n-dimensional space, wherein the computer compares in the n-dimensional space the received phoneme with each phoneme in the database of stored phonemes.

31. The system of claim 30, wherein the computer recognizes the received phoneme using the comparison in the n-dimensional space of the received phoneme with each phoneme in the database of stored phonemes.

32. The system of claim 31, wherein the computer compares the received phoneme with each phoneme in the database of stored phonemes by comparing a first distance from a center of the n-dimensional space to a first point associated with the received phoneme with a second distance from the center of the n-dimensional space to a second point associated with each respective stored phoneme from the database of stored phonemes.

33. The system of claim 32, wherein the computer recognizes the received phoneme by determining a difference between the first distance and the second distance.

34. The system of claim 33, wherein the computer recognizes the received phoneme as associated with a stored phoneme corresponding to a shortest distance between the first distance and the second distance.

35. A medium storing a program for instructing a computer device to recognize a received speech signal using a database of stored phonemes converted into n-

dimensional space, the program comprising instructing the computer device to perform the following steps:

receiving a received phoneme;

converting the received phoneme to n-dimensional space;

comparing the received phoneme to each of the stored phonemes in n-dimensional space; and

recognizing the received phoneme according the comparison of the received phoneme to each of the stored phonemes.

36. A medium storing a program for instructing a computer device to recognize a received speech signal using a database of stored phonemes converted into n-dimensional space, the database of stored phonemes formed by training the stored phonemes according to the following steps:

(1) determining a phoneme vector as a time-frequency representation of the stored phoneme;

(2) dividing the phoneme vector into phoneme segments;

(3) assigning each phoneme segment into a plurality of phoneme parameters;

(4) expanding each phoneme segment and plurality of phoneme parameters into an expanded stored-phoneme vector with expanded vector parameters;

(5) transforming the expanded stored-phoneme vector into an orthogonal form using singular-value decomposition wherein:

$[x_1 x_2 \ldots x_m] = [u_1 u_2 \ldots u_m] \Lambda V^t$, where x_k is a k^{th} acoustic vector for a corresponding stored phoneme, u_k is the corresponding orthogonal vector and Λ and V are diagonal and unitary matrices, respectively, the program stored on the medium instructing the computer device to perform the following steps:

(1) receiving an analog acoustic signal;

(2) converting the analog acoustic signal into a digital signal;

(3) determining a received-signal vector as a time-frequency representation of the received digital signal;

(4) dividing the received-signal vector into received-signal segments;

(5) assigning each received-signal segment into a plurality of received-signal parameters;

(6) expanding each received-signal segment and plurality of received-signal parameters into an expanded received-signal vector;

(7) transforming the expanded received-signal vector into an orthogonal form using singular-value decomposition wherein:

$[y_k] = [z_k] \Lambda V^t$, where y_k is a k^{th} acoustic vector for a corresponding received phoneme, z_k is the corresponding orthogonal vector and Λ and V are diagonal and unitary matrices, respectively;

(8) determining a first distance associated with the orthogonal form of the expanded received-signal vector and a second distance associated respectively with each orthogonal form of the expanded stored-phoneme vectors; and

(9) recognizing the received phoneme according to a comparison of the first distance with the second distance.